Robust optimization in simulation

Dellino, G.; Kleijnen, Jack P.C.; Meloni, C.

Published in:
International Journal of Production Economics

Publication date:
2010

Citation for published version (APA):
Dellino, G., Kleijnen, J. P. C., & Meloni, C. (2010). Robust optimization in simulation: Taguchi and Response Surface Methodology. *International Journal of Production Economics*, *125*(1), 52-59. http://www.sciencedirect.com/science/article/pii/S0925527309004502

Robust optimization in simulation: Taguchi and Response Surface Methodology Online companion paper

Gabriella Dellino [a)*], Jack P.C. Kleijnen [a)], and Carlo Meloni [b)]

[a)] *Department of Information Management /CentER, Faculty of Economics and Business Administration, Tilburg University, Postbox 90153, 5000 LE Tilburg, Netherlands, e-mail:G.Dellino@uvt.nl*
[*] *Corresponding author*
[a)] *Department of Information Management /CentER, Faculty of Economics and Business Administration, Tilburg University, Postbox 90153, 5000 LE Tilburg, Netherlands, http://center.uvt.nl/staff/kleijnen/*
[b)] *Department of Electrical Engineering and Electronics, Polytechnic of Bari, Via E. Orabona 4, 70125 Bari, Italy, e-mail:meloni@deemail.poliba.it*

25 August 2008

Abstract

This online companion paper supplements the short version (with the same title) that was submitted for publication. It discusses robust optimization in a wider context, including additional references. It details Taguchi's world view and RSM for robust optimization. Furthermore it presents results for additional experiments with a smaller range of the order quantity, and a smaller value for the demand variance. The bootstrap results are verified through replication of the whole experiment, so-called macroreplication.

1 Introduction

This online companion paper add details to the short version that is also titled 'Robust optimization in simulation: Taguchi and Response Surface Methodology'. We try to minimize the overlap between the two versions, while maintaining the flow of thought in this online paper. Even though we suggest the reader to use this companion paper as an addendum, we make this paper—to a certain extent—readable on its own; e.g., we add some references that are also given in the short version (but the online paper has many new references).

Furthermore, each section of this online paper corresponds with the section with the same title in the short version; e.g., Section 1 adds details to the section with the same title in the short version. The long Subsection 4.2, however, is split into subsubsections to improve the readability of that subsection.

The importance of *optimizing* engineered systems (artifacts) is also emphasized in the 2006 NSF panel reported in Oden (2006). That report also points out the crucial role of *simulation* in engineering science. The simulation model may be either deterministic or random. We, however, focus on deterministic simulation. Nevertheless, we expect that our new methodology can also be applied to find the optimal inputs for random simulation models and real-world systems.

In practice, some inputs of the given simulation model are uncertain so the optimum solution that is derived—ignoring these uncertainties—may be completely wrong. In a different context—namely Linear Programming (LP)—Ben-Tal mentions that 13 of the approximately 100 LP models in the NETLIB Library give constraint violations (infeasibility) when perturbing the input data by only 0.01% (see also Ben-Tal and Nemirovski (2008)). Simulation models are more difficult compared with LP and Non-Linear Programming (NLP) models:

- Simulation models treated as black boxes imply *implicit* functions for the goal and constrained outputs.
- Simulation models are *dynamic* (whereas LP and NLP models are usually static).

A well-known distinction in the management literature (see the many references in Kleijnen (1980)) is

- Operational decisions: repetitive decisions (e.g., daily inventory management)
- Strategic decisions: one-shot decisions (e.g., designing a computerized inventory management system).

We focus on strategic decisions (for operational decisions, Control Theory seems more appropriate). These decisions may concern the design of either products or processes (for manufacturing these products). *Robust design* is important for engineers, in many disciplines. Actually, these engineers should work together, which results in *Multidisciplinary Design Optimization* (MDO); see Alexandrov and Hussaini (1997) and Beyer and Sendhoff (2007). Products of *Computer Aided Design* (CAD) and *Computer Aided Engineering* (CAE) are airplanes, automobiles, TV sets, chemical plants, computer chips, etc.—developed at companies such as Boeing, General Motors, and Philips. Recent surveys are Chen et al. (2003), Chen et al. (2006), Meckesheimer et al. (2001), Oden (2006), and Simpson et al. (2001).

The literature (see Beyer and Sendhoff (2007) and Kleijnen (2008)) distinguishes the following two approaches to strategic decision-making in an uncertain world (Park et al. (2006) also detail the first approach, and discuss more approaches):

- *Taguchi*'s approach, originally developed to help Toyota design 'robust' cars; i.e., cars that perform reasonably well in many circumstances (from the snows in Alaska to the sands in the Sahara). Taguchi is a Japanese engineer and statistician; see Taguchi (1987) and Wu and Hamada (2000).
- *Robust Optimization* (RO)—developed by Ben-Tal, Nemirovsky, Bertsimas and others—to make the original Mathematical Programming (MP) solutions less sensitive to perturbations in the coefficients of the MP models; see Ben-Tal and Nemirovski (2008) and also Beyer and Sendhoff (2007) and Greenberg and Morrison (2008). (Stochastic MP is a related yet different approach; see Mulvey, Vanderbei, and Zenios (1995) and also Beyer and Sendhoff (2007), Greenberg and Morrison (2008), and Sahinidis (2004).)

In practice, classic (standard) optimization may be counterproductive! Indeed, the French say (in translation): 'the best is the enemy of the better'; and Marczyk (2000, p. 3) states: 'Optimization is actually just the opposite of robustness'.

The rest of this article is organized as follows. Section 2 details Taguchi's worldview. Section 3 details RSM for robust optimization. Section 4.1 presents results for an experiment with a smaller range of the order quantity. Section 4.2 verifies the bootstrap results through the use of macroreplicates. Appendices gives technical details. An extensive list of references enables the reader to study robust simulation-optimization in more detail.

2 Taguchi's worldview

Figure 1 illustrates Taguchi's distinction between two types of variables:

- Decision (or control) factors, which we denote by d_j ($j = 1, \ldots, k$).
- Environmental (or noise) factors, which we denote by e_g ($g = 1, \ldots, c$).

Other authors distinguish between environmental uncertainty (e.g., demand uncertainty) and system uncertainty (e.g., yield uncertainty); see Mula et al. (2006) and also Beyer and Sendhoff (2007). Implementation errors may also be a source of uncertainty. These errors occur whenever recommended (optimal) values of control factors are to be realized in practice; see Stinstra and Den Hertog (2007). Continuous values are hard to realize in practice, because only limited accuracy is then possible; e.g., the EOQ turns out to be the square

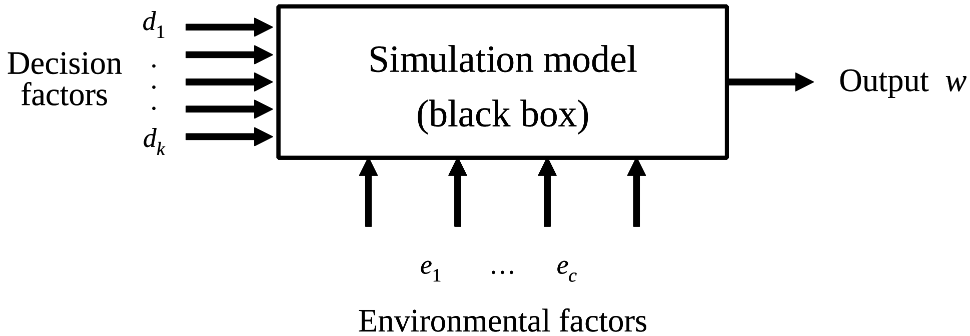

Fig. 1. Taguchi's view

root of some expression, but in practice only a discrete number of units can be ordered. Besides implementation errors, there are validation errors of the simulation model (compared with the real system) and the metamodel (compared with the simulation model); see Kleijnen and Sargent (2000).

The relevant problem formulation depends on the risk attitude of the users (they might be risk-seeking optimists), which may vary with the application. We conjecture that our heuristic also applies to alternative problem formulations, but in this article we do not investigate these alternatives. Many references on supply-chain risk-management are given in Wu et al. (2008), who focus on the mean-variance trade-off in the newsvendor's inventory problem. The mean-variance trade-off for simulation models is also examined by Apley, Liu, and Chen (2006) and Chen, Jin, and Sudjianto (2006).

3 RSM and robust optimization

It is convenient and traditional in Design Of Experiments (DOE) to use *coded*—also called *standardized* or *scaled*—factor values. Let the experiment consist of n factor combinations of the 'original' factors z_j that corresponds with d_j or e_g; furthermore, let l_j denote the lowest value of z_j in the experiment, and u_j the highest ('upper') value. Then the coded variable x_j use the linear transformation

$$x_j = a_j + b_j z_j \text{ with } a_j = \frac{l_j + u_j}{l_j - u_j} \text{ and } b_j = \frac{2}{u_j - l_j}. \tag{1}$$

The term $(u_j - l_j)$ is the *range* of input j. If z is a random variable (like e), then this coding implies $var(x) = b^2 var(e)$. The numerical accuracy of the estimates may be affected by coding; we focus on the estimated effects of the coded variables. Coding is further discussed by Kleijnen (2008, p. 29).

Myers and Montgomery (1995, p. 493-494) assume that the environmental variables $\mathbf{e}$ satisfy

$$E(\mathbf{e}) = \mathbf{0} \text{ and } \mathbf{cov}(\mathbf{e}) = \sigma_e^2 \mathbf{I}, \tag{2}$$

and derive the mean and the variance of y (the regression predictor of the simulation output w), after averaging over the noise factors:

$$E(y) = \beta_0 + \boldsymbol{\beta}'\mathbf{d} + \mathbf{d}'\mathbf{B}\mathbf{d} \tag{3}$$

and

$$var(y) = \sigma_e^2(\boldsymbol{\gamma}' + \mathbf{d}'\boldsymbol{\Delta})(\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{d}) + \sigma_\epsilon^2 = \sigma_e^2 \mathbf{l}'\mathbf{l} + \sigma_\epsilon^2, \tag{4}$$

where $\mathbf{l} = (\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{d}) = (\partial y/\partial e_1, \ldots, \partial y/\partial e_c)'$; i.e., $\mathbf{l}$ is the gradient with respect to the environmental factors. So, the larger the gradient's elements are, the larger the variance of the predicted simulation output is—which stands to reason. Furthermore, if $\boldsymbol{\Delta} = \mathbf{0}$ (no control-by-noise interactions), then $var(y)$ cannot be controlled through the control variables $\mathbf{d}$.

Equation (4) implies that the predicted simulation output y has heterogeneous variances—even if σ_e^2 and σ_ϵ^2 were constants—because changing the control factors $\mathbf{d}$ changes $var(y)$. Whereas Myers and Montgomery (1995) present examples with $\sigma_e^2 = \sigma_\epsilon^2/2$, Kleijnen (2008, p. 136) gives a supply-chain simulation with $\sigma_e^2 = 10\sigma_\epsilon^2$. Most important is the gradient $\mathbf{l}$, because it shows the key role played by the control-by-noise interactions; i.e., to reduce the predicted output's variance $var(y)$(or σ_y^2) the analysts should take advantage of the interactions $\boldsymbol{\Delta}$; they cannot control the main effects of the noise factors ($\boldsymbol{\gamma}$) and the variances of the noise factors and the residuals (σ_e^2 and σ_ϵ^2). For example, if a particular decision factor (say, d_1) has no effects on the mean output (so $\beta_1 = \beta_{1;1} = \beta_{1;2} = \ldots = \beta_{1;k} = 0$) but has important interactions with the noise factors (e.g., $\delta_{1;2} >> 0$), then this interaction can be utilized to decrease the output variance (e.g., decrease σ_y^2 by decreasing d_1). If there are multiple decision factors, then the following solution method may be tried:

(1) select the values of some decision factors such that $\mathbf{l} = \mathbf{0}$, so $var(y)$ in (4) is minimized;

(2) select the remaining decision factors such that the predicted mean output $E(y)$ in (3) gets the desired value.

Notice that we might try to find a design that is 'D-optimal'; i.e., a design that minimizes the determinant of $\mathbf{cov}(\widehat{\boldsymbol{\zeta}})$; see Chung, Goldfarb, and Montgomery (2007).

In order to apply classic OLS results, we assume that σ_w^2 is constant, the outputs for different scenarios are independent, and the environmental factors are fixed (Myers and Montgomery (1995, p. 490) do not make these assumptions explicit; in our EOQ application we can derive the true Pareto optimum, so we can verify how sensitive our analysis is to these assumptions).

The t statistic is used to test the importance of an individual factor effect.

It is well-known that the t statistic is not very sensitive to nonnormality; see Kleijnen (1987). The 'reduced' metamodel (nonsignificant effects eliminated) may imply a unique optimum, whereas the full metamodel may suggest (say) a saddlepoint. To find the unimportant effects, Myers and Montgomery (1995, p. 487) use ANalysis Of VAriance (ANOVA). Note that $t_{n-q}^2 = F_{1;n-q}$; the F statistic is used in ANOVA.

Notice that Myers and Montgomery (1995, p. 508) discuss the use of transformations of the dependent variable, before performing the regression analysis; also see Kleijnen (2008, p. 98).

To test the validity of the RSM metamodel, the linear regression literature offers several methods. We focus on a method that is also applied outside linear regression (e.g. in Kriging), namely cross-validation. There are several variations on cross-validation (see Iooss, Ribatet, and Marrel (2007) and Meckesheimer et al. (2001)), but the most popular variant is *leave-one-out cross-validation*.

4 EOQ inventory simulation

4.1 Simulation optimization of the EOQ model

We program the simulation model in Arena; see Kelton, Sadowski, and Sturrock (2007). After running this simulation, we also compute the estimated effects of the *original* (not standardized) variable; e.g., the estimated quadratic effect is then of order 10^{-6}, so it seems unimportant; however, Q^2 is $30298^2 = 9 \times 10^8$ so their joint effect is of order 10^2.

We also experiment with a *smaller* experimental area; i.e., a smaller Q range. We assume that the center of this new area is still close to the true optimum. The Taylor series argument suggests that this smaller area gives a better approximation locally. Appendix 1 shows that the smaller Q range indeed gives a more accurate metamodel; the resulting estimated optimum is only 1% below the true EOQ and the corresponding cost virtually equals the true cost.

4.2 Robust optimization of EOQ model

Notice that Yu (1997) also assumes an uncertain demand rate, but uses other criteria than we do: he either minimizes the maximum costs or minimizes the maximum percentage deviation from the optimal cost. Moreover he does not assume a probability function for the various scenarios (demand rate values), but uses 'a discrete scenario set'. Altogether, his approach resembles that of Ben-Tal et al., which we discussed in the Introduction.

Notice further that the assumption of uncertain constants is often made in deterministic simulation of physical systems; e.g., a nuclear waste-disposal simulation may assume that the permeability of a specific area is constant but unknown; see Kleijnen and Helton (1999). An economic example is the exchange rate between the US dollar and the euro exactly one year from today: that rate is a constant but unknown.

We may collect historical data to infer the probability of the true value of the parameter a. If there is no such data, then we may ask experts for their opinion on the true value of the parameter. This *knowledge elicitation* results in an input distribution (say) $F(a)$. In practice, several distribution types are used, such as normal, lognormal, and uniform; see Kleijnen and Helton (1999). In our experiments we assume $a \sim N(\mu_a, \sigma_a)$.

4.2.1 LHS design

LHS splits the range of possible a values ($0 < a < \infty$) into $n_e = 5$ equally likely subranges, namely $(0, \mu_a - 0.85\sigma_a]$, $(\mu_a - 0.85\sigma_a, \mu_a - 0.73\sigma_a]$, $(\mu_a - 0.73\sigma_a, \mu_a + 0.73\sigma_a]$, $(\mu_a + 0.73\sigma_a, \mu_a + 0.85\sigma_a]$, $(\mu_a + 0.85\sigma_a, \infty)$. Notice that the 'base' value μ_a has zero probability, but a value 'close' (namely less than $0.73\sigma_a$ away) has 20% probability. Besides the relatively high uncertainty $\sigma_a = 0.50\mu_a$, Appendix 2 shows the results for the smaller uncertainty $\sigma_a = 0.10\mu_a$.

We again code the inputs; see (1). So x_1 corresponds with Q and x_2 with a; e.g., $a = 7687{,}37$ corresponds with $x_2 = -0.1017$ (not exactly zero, because of the sampling that LHS does). Furthermore, if $\sigma_a = 0.50\mu_a = 4000$ and $b_2 = 2.85 \times 10^{-4}$, then the standard deviation of x_2 is $\sigma_2 = 4000 \times 2.85 \times 10^{-4} = 1.14$.

This gives the estimated effects displayed in the row denoted by 0 (zero rows eliminated) in Table 1. The rest of this table displays the cross-validation results. This table gives the scatterplot in Figure 2. This table and this figure suggest that this metamodel is adequate for robust optimization through RSM. Comparing the plots for low and high uncertainties suggests that the first plot is much worse; however, using the same scale in both plots (not displayed) changes that impression.

Appendices 1 and 2 give results for a smaller range of the decision variable Q and the environmental variable a. These results show even better fit.

Notice that σ_a^2 is a known input value, so we also know the variance of the corresponding coded variable x_2, namely $\sigma_2^2 = 1.14^2 = 1.3$). Altogether we

i	$\widehat{\beta_{0(-i)}}$	$\widehat{\beta_{1(-i)}}$	$\widehat{\beta_{1;1(-i)}}$	$\widehat{\gamma_{1(-i)}}$	$\widehat{\delta_{1;1(-i)}}$	$\widehat{y_{(-i)}}$	$\widehat{y_{(-i)}}/C_i$
0	88150.40	190.56	1058.33	36774.03	-899.67		
1	88144.21	172.94	1088.31	36755.96	-863.54	51440.09	1.005
2	88147.70	181.73	1072.54	36768.01	-887.64	61545.93	1.002
3	88152.19	198.89	1046.41	36774.09	-899.80	85169.34	0.999
4	88154.29	214.81	1026.29	36764.26	-880.14	102725.72	0.997
5	88157.15	259.16	976.22	36714.38	-780.37	126368.95	0.994
6	88150.48	190.51	1058.24	36773.93	-899.57	51096.08	1.000
7	88154.53	188.27	1054.63	36770.90	-896.54	61150.15	1.001
8	88164.19	182.52	1046.82	36773.90	-899.54	84550.05	1.002
9	88172.91	177.00	1040.41	36784.95	-910.59	101956.72	1.003
10	88190.37	165.55	1028.40	36817.51	-943.16	125653.78	1.004
11	88124.57	190.56	1090.86	36793.63	-899.67	51330.94	0.994
12	88131.43	190.56	1081.72	36783.93	-899.67	61275.30	0.997
13	88146.40	190.56	1063.03	36774.05	-899.67	84407.52	1.000
14	88158.08	190.56	1049.58	36776.69	-899.67	101600.85	1.001
15	88177.60	190.56	1028.69	36795.56	-899.67	124973.16	1.002
16	88136.05	182.81	1071.52	36789.93	-883.77	52147.29	0.996
17	88137.51	183.42	1069.82	36783.76	-889.94	61965.54	0.997
18	88139.92	184.45	1067.07	36774.13	-899.57	84805.76	0.998
19	88141.21	185.03	1065.63	36769.57	-904.12	101775.07	0.999
20	88142.70	185.75	1064.09	36765.65	-908.04	124813.22	0.999
21	88140.63	218.39	1105.69	36745.49	-956.75	53675.97	1.008
22	88144.24	210.75	1090.83	36760.27	-927.18	63283.89	1.005
23	88148.76	198.15	1069.18	36773.97	-899.79	85768.71	1.001
24	88150.72	188.53	1055.65	36773.21	-901.30	102506.95	1.000
25	88152.94	164.73	1027.41	36751.56	-944.60	125152.04	0.998

Table 1
Cross-validation of regression metamodel for RO of EOQ

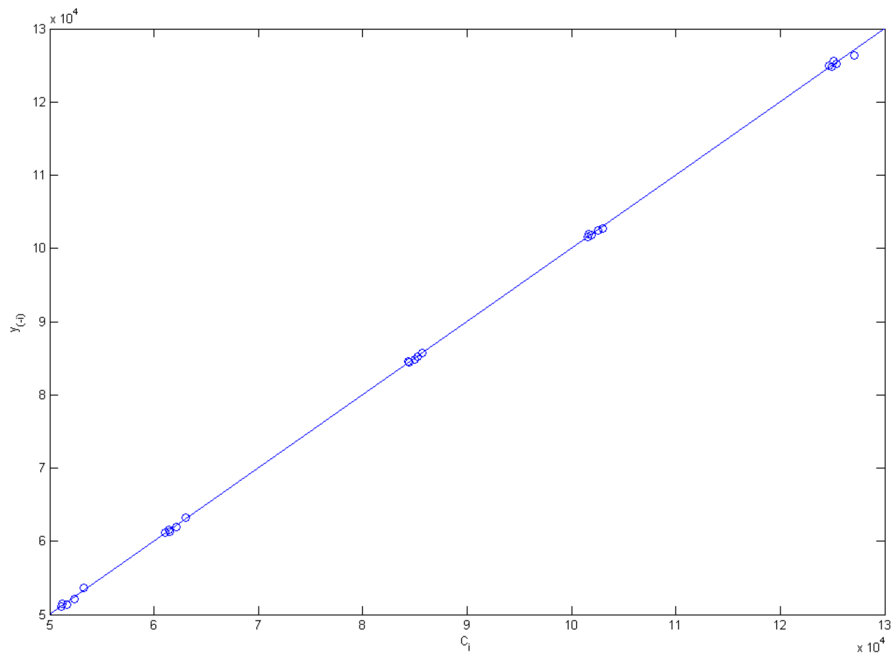

Fig. 2. Scatterplot of regression metamodel for RO of EOQ

obtain $\widehat{\sigma_C} = [(\widehat{\gamma_1}+\widehat{\delta_{1;1}}x_1)^2\sigma_2^2+\widehat{\sigma}_\epsilon^2]^{1/2} = [(36755.96-863.54x_1)^2 \times 1.3+4.6224\times 10^4]^{1/2}$.

We repeat the experiment with a smaller σ_a The new threshold values give the estimated Pareto frontier of Figure 3. Comparing the estimated Pareto frontier for this small σ_a (reported in this companion paper) and for high σ_a (reported in the short paper) demonstrates that a less volatile world gives lower mean cost.

Notice that we focus on estimating the variability of the Pareto curve, but we could also have estimated the variability of the solution of the robust optimum problem. So the B bootstrap regression parameters ζ^* gives B values for Q^+ and the corresponding C^+ and $s(C)^+$. These B values can be used to derive a CI; see Efron and Tibshirani (1993).

4.2.2 Macroreplicates

Actually, we can validate our (fast) bootstrap procedure as follows Our EOQ simulation is the opposite of *expensive simulation*: some realistic simulations take hours or weeks for a single run, whereas bootstrapping this simulation's results still takes only seconds. So we repeat our LHS sample (say) L times; i.e., we sample the demand rate a from the normal distribution cut-off at zero, while keeping the five Q values fixed. This sample of L *macroreplicates* gives

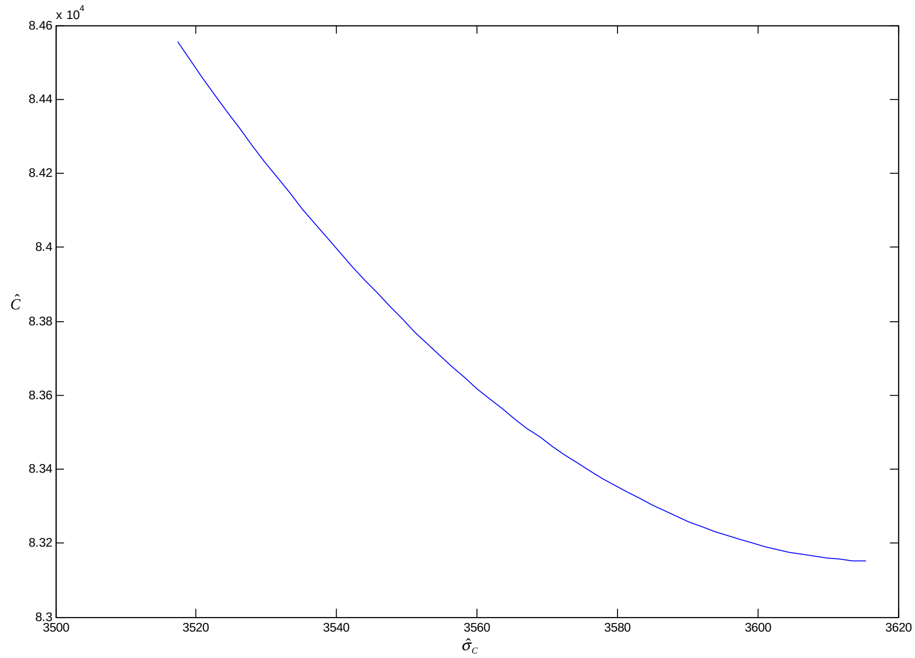

Fig. 3. Less volatile world: estimated Pareto frontier for EOQ simulation with threshold

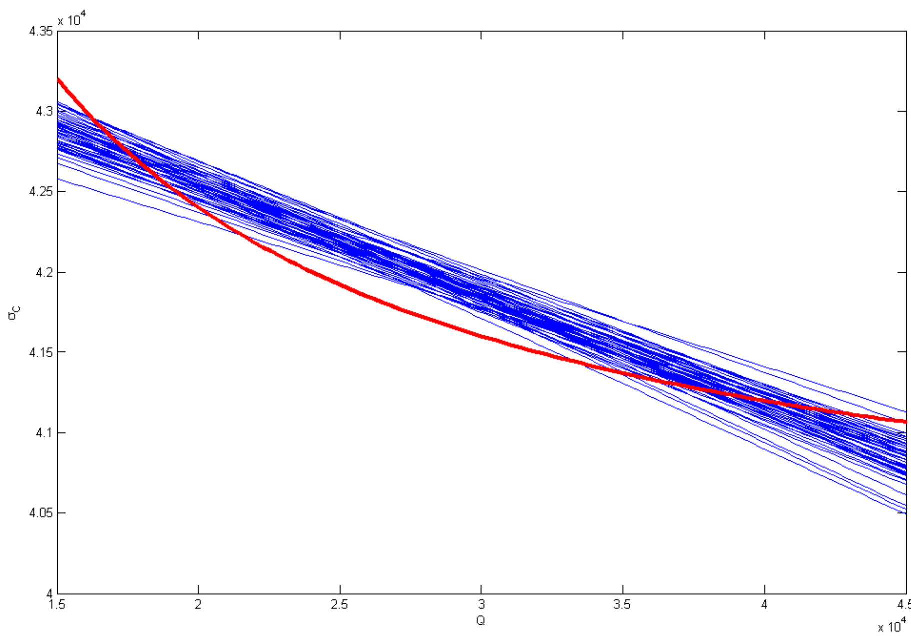

Fig. 4. Bootstrapped standard deviations of the cost, and true standard deviation of the cost (heavy curve)

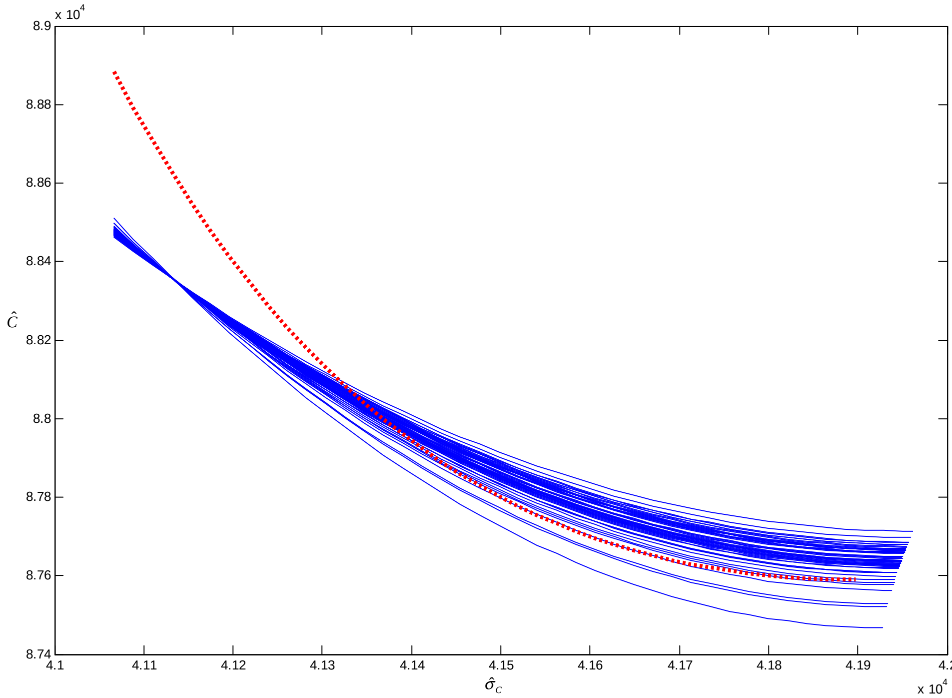

Fig. 5. Pareto frontiers estimated from 50 macroreplicates, and true frontier (dotted curve)

the regression estimate $\widehat{\boldsymbol{\zeta}}_l$ with $l = 1, \ldots, L$. This $\widehat{\boldsymbol{\zeta}}_l$ gives $\widehat{C_l}$ (costs estimated through RSM metamodel) and $\widehat{\sigma_{C_l}}$ (corresponding standard deviation). Together with the threshold T this gives the estimated Pareto frontier. Repeating this LHS L times gives a set of L estimated Pareto frontiers; see Figure 5 with $L = 50$. This figure suggests that these estimated curves all intersect near the point $(4.11, 8.83)$, but zooming-in around this point reveals that the 50 curves do not intersect in a single point. Appendix 3 also displays the 50 $\widehat{C}$-curves and the 50 $\widehat{\sigma_C}$-curves. These curves results in 50 Pareto curves estimated from 50 macroreplicates; see again Figure 5. This figure assumes that a second-order polynomial is a perfect approximation of the true I/O function, whereas the true EOQ formulas show that this assumption is false This figure shows that the macroreplicates give a tighter bundle than bootstrapping does. Appendix 3 shows that this phenomenon is explained by the negative correlations between estimated regression coefficients in the macroreplicates. In general, we could argue that —compared with bootstrapping—macroreplicates use much more computer time, and provide more information so the spread in the estimated Pareto curves is smaller. Appendix 3 also shows that if we replace LHS by crude sampling in the macroreplicates, then a bigger spread is the result; i.e., LHS is indeed a variance reduction technique.

Q	a	C
19393.40	5559.67	61945.81
19393.40	10529.69	114721.4
40606.60	5559.67	63330.64
40606.60	10529.69	114499.6
15000	8044.68	89132.55
45000	8044.68	89342.05
30000	4530.34	51615.54
30000	11559.02	124713.8
30000	8044.68	88164.67

Table 2
I/O simulation data for EOQ model with CCD design

4.2.3 CCD

Finally, we compare the (traditional Taguchian) crossed design (used in the short paper) with a CCD. A CCD for two factors (Q and a) consists of a 2^2 design (the four combinations of the two extreme values per factor -1 and 1), the four 'axial' points $((0,-\sqrt{2})$, $(0,\sqrt{2})$, $(-\sqrt{2},0)$, $(\sqrt{2},0))$, and the central point $((0,0))$ in coded values; the value $\sqrt{2}$ is selected to make the CCD 'rotatable' (see Myers and Montgomery (1995, p. 299)) The original input values plus the corresponding output values are displayed in Table 2.

Note: A CCD is not a subset of the crossed design, because a CCD does not sample any factor value, whereas the crossed design uses LHS for the environmental factor a. Consequently, that crossed design does not have (say) coded values -1 and 1 for a, which are at exactly the same distance from 0.

We again validate the resulting metamodel through cross-validation; see Appendix 4 for details. We repeat our analysis for this CCD. This gives the Pareto frontier of Figure 6. So the CCD with its nine combinations gives a better estimate of the true frontier than the 5×5 crossed-design does.

Appendix 1: Smaller Q-range

Row 1 of Table 3 shows the Q values in the smaller experimental area; row 2 gives the corresponding simulation outputs.

Regression analysis of the I/O data in Table 3 gives Table 4 and the scatterplot of Figure 7. This table and this scatter plot imply that the smaller Q-range gives a more accurate metamodel. The new estimated optimum $\widehat{Q}_o$ is 25115, which gives $\widehat{C}_o = 87607$ so $\widehat{Q}_o/Q_o = 25115/25298 = 0.99$ and $\widehat{C}_o/C_o$

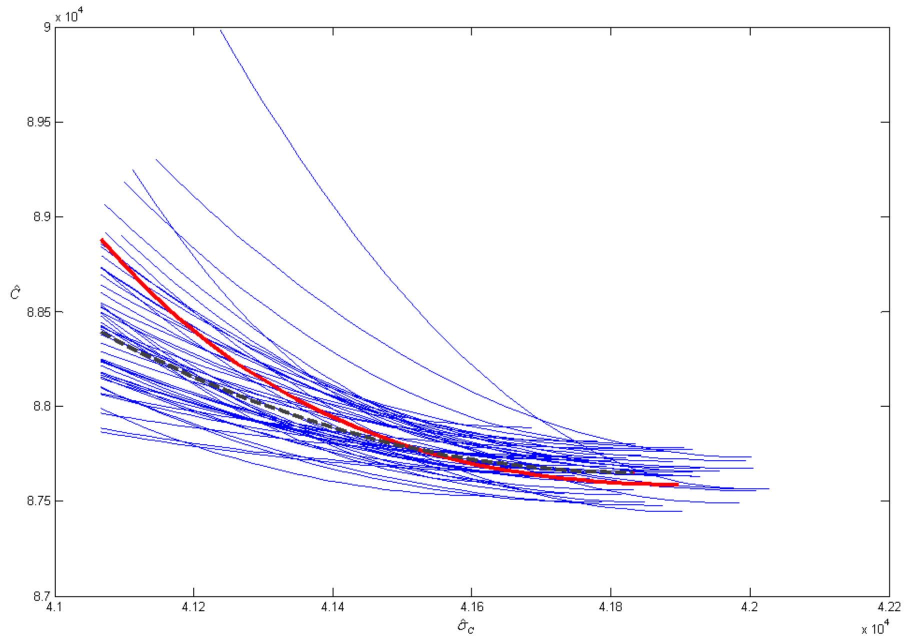

Fig. 6. Bootstrapped Pareto frontiers, original estimated frontier (dashed curve) and true Pareto frontier (heavy curve) based on CCD

Q	22500	26250	30000	33750	37500
C	87641.66	87594.64	87700	87906.95	88185

Table 3
I/O data for EOQ simulation with smaller experimental area

i	$\widehat{\beta_{0(-i)}}$	$\widehat{\beta_{1(-i)}}$	$\widehat{\beta_{1;1(-i)}}$	$\widehat{y_{(-i)}}$	$\widehat{y_{(-i)}}/C_i$
0	87698.26	279.798	214.78		
1	87704.57	309.26	172.69	87633.242	0,999
2	87707.76	274.26	206.86	87612.056	1,000
3	87696.62	279.798	216.71	87698.26	0,9999
4	87690.03	274.99	221.64	87891.854	0,9997
5	87692.38	307.23	253.97	88192.838	1,001

Table 4
Cross-validation of EOQ regression metamodel with smaller range

$= 87618/87589 = 1.0003$. Comparison with the old results ($\widehat{Q_o}/Q_o = 1.13$ and $\widehat{C_o}/C_o = 1.001$) shows that the smaller Q range improves the estimated optimum.

Appendix 2: Smaller uncertainty $\sigma_a = 0.10\mu_a$

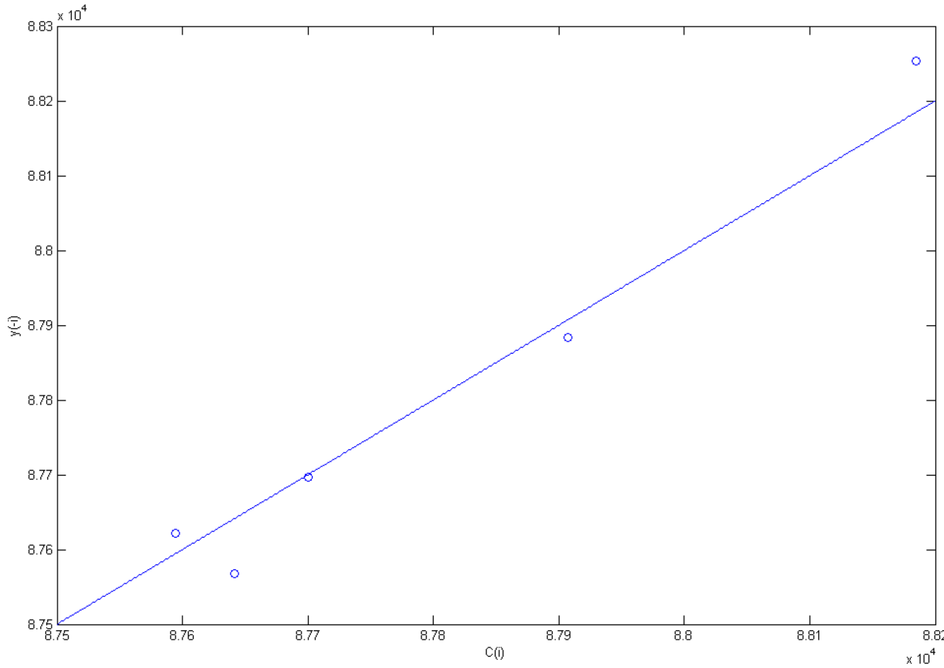

Fig. 7. Scatterplot of the EOQ regression metamodel for smaller Q-range

$Q \setminus a$	6076.55	7438.96	7.832.04	8595.36	9101.10
15000	67876.77	82590.79	86836.02	95079.90	111541.90
22500	67381.35	81732.06	85872.48	93912.80	99239.95
30000	67696.14	81865.20	85953.20	93891.76	99151.47
37500	68335.02	82395.09	86451.64	94329.13	99548.38
45000	69135.94	83123.34	87158.94	94995.71	100188

Table 5
I/O simulation data for EOQ model with smaller a-range

Tables 5 and 6 together with Figure 8 give results for smaller uncertainty in the demand rate.

Appendix 3: Macroreplicates

Figures 9 and 10 display the 50 $\widehat{C}$-curves and the 50 $\widehat{\sigma_C}$-curves respectively, computed from 50 macroreplicates. Note that the latter figure suggests that the 50 estimated curves coincide, but zooming-in reveals that the 50 curves do not coincide: these curves have little spread; see Figures 11 and 12. We point out that each macroreplicate gives a different mean and standard deviation for the coded variable x_2; e.g., $x_{2;l} = min_k a_{l;k}$ with $l = 1, \ldots, 50$ and $k = 1, \ldots, 5$.

i	$\widehat{\beta_{0(-i)}}$	$\widehat{\beta_{1(-i)}}$	$\widehat{\beta_{1;1(-i)}}$	$\widehat{\gamma_{1(-i)}}$	$\widehat{\delta_{1;1(-i)}}$	$\widehat{y_{(-i)}}$	$\widehat{y_{(-i)}}/C_i$
0	83374.03	307.26	1070.93	15824.45	-387.14		
1	83373.72	296.40	1082.41	15814.75	-367.74	67977.23	1.001
2	83375.07	313.61	1062.48	15825.97	-390.18	82516.93	0.999
3	83375.95	316.08	1058.27	15824.30	-386.85	86725.30	0.999
4	83378.77	320.83	1047.87	15815.61	-369.45	94878.21	0.998
5	83382.36	325.08	1036.43	15802.71	-343.66	100240.08	0.997
6	83383.68	300.95	1064.25	15813.17	-375.86	67498.17	1.002
7	83387.18	299.36	1060.43	15820.66	-383.35	81915.79	1.002
8	83388.19	299.02	1059.09	15824.72	-387.41	86079.62	1.002
9	83390.53	298.34	1055.76	15836.08	-398.77	94178.12	1.003
10	83392.58	297.83	1052.67	15847.45	-410.14	99559.36	1.003
11	83353.46	307.26	1092.07	15842.31	-387.14	67511.15	0.997
12	83367.41	307.26	1078.48	15825.81	-387.14	81799.08	0.999
13	83370.36	307.26	1075.25	15824.40	-387.14	85915.31	1.000
14	83375.54	307.26	1069.02	15825.18	-387.14	93908.39	1.000
15	83378.87	307.26	1064.54	15828.47	-387.14	99207.34	1.001
16	83355.49	295.13	1083.75	15846.11	-365.47	68110.62	0.997
17	83362.44	300.30	1080.18	15827.79	-383.80	82233.12	0.998
18	83363.51	301.14	1079.72	15824.25	-387.34	86297.79	0.998
19	83364.95	302.35	1079.27	15818.05	-393.54	94183.08	0.998
20	83365.52	302.94	1079.30	15813.89	-397.69	99401.86	0.999
21	83372.90	346.65	1112.57	15789.26	-457.51	69500.37	1.005
22	83372.29	317.81	1084.94	15821.92	-392.19	83245.96	1.001
23	83372.43	314.61	1081.48	15824.57	-386.90	87251.28	1.001
24	83373.06	310.04	1075.65	15826.26	-383.51	95037.07	1.000
25	83373.96	307.40	1071.20	15824.62	-386.79	100190.39	1.000

Table 6
Cross-validation of regression metamodel for RO of EOQ with smaller a-range

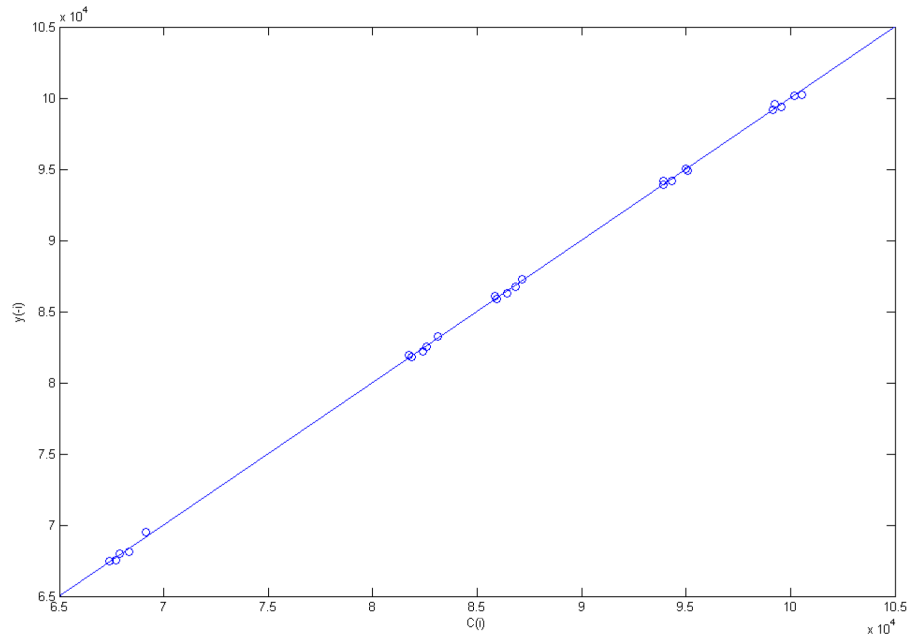

Fig. 8. Scatterplot of the EOQ regression metamodel for smaller a-range

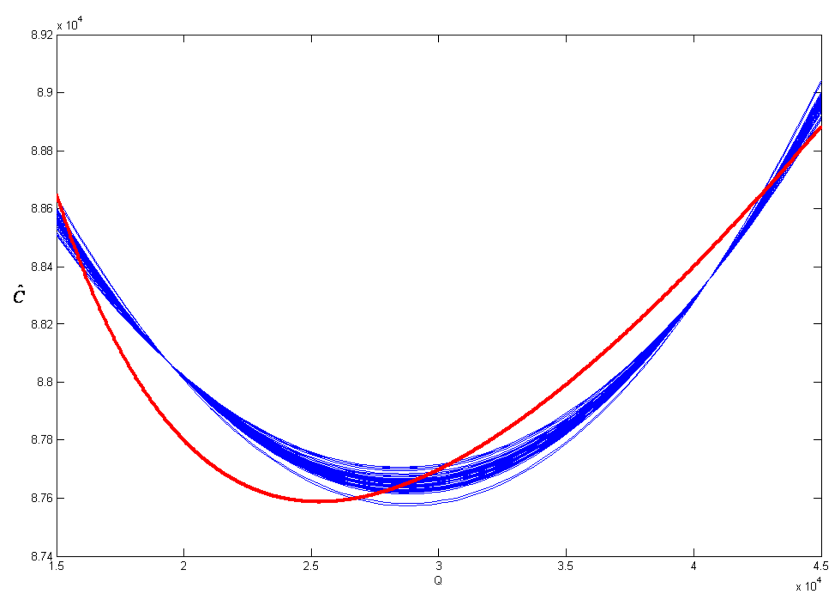

Fig. 9. Replicated estimated costs, and true cost (heavy curve)

There is no solution for the constrained optimization problem if the LHS happens to result in an extremely high $\widehat{\sigma}(C)$. Actually this happened once in our 50 macroreplicates; we simply threw away this macroreplicate, and sampled again.

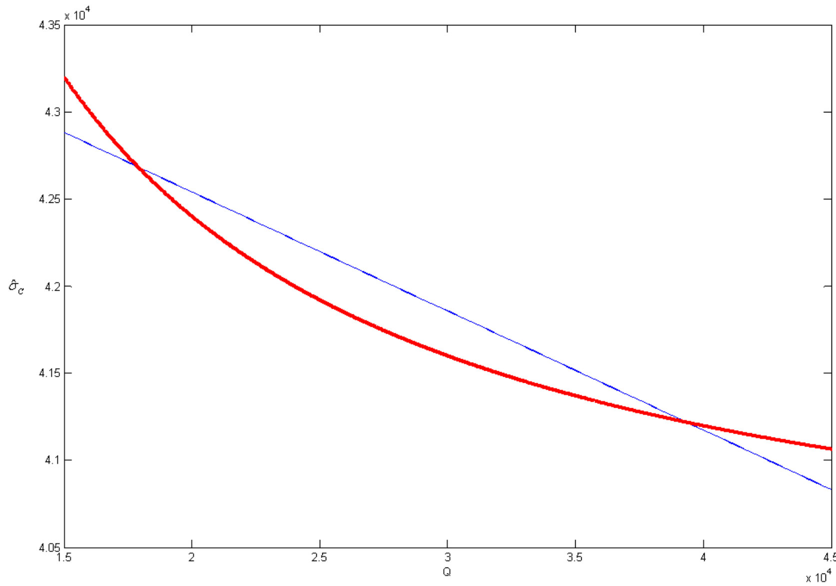

Fig. 10. Replicated standard deviations of the cost, and true standard deviation (heavy curve)

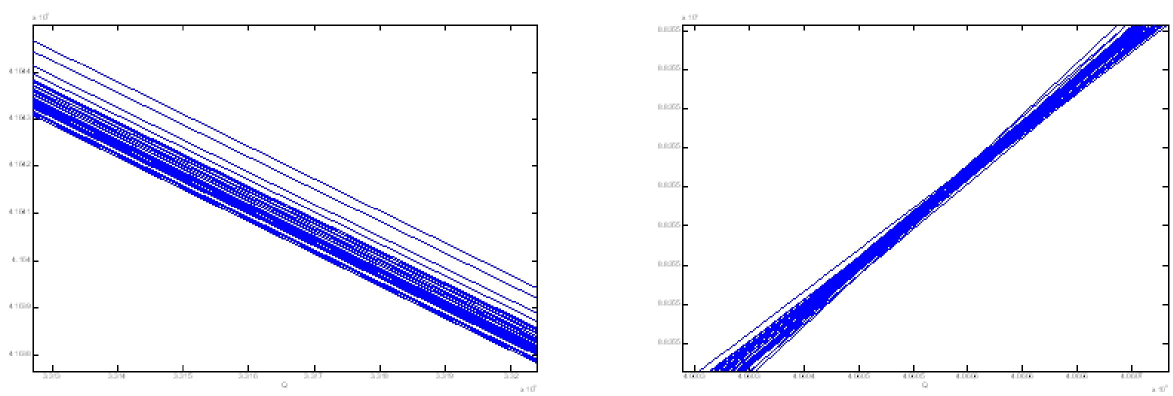

Fig. 11. Zoom: Mean Cost, estimated through 50 macroreplicates

Figure 13 shows that if we replace LHS by crude sampling in the macroreplicates, then bigger spread results. This bigger spread is caused by a bigger spread in the estimated regression coefficients; e.g. Figure 14 shows the Box plot for the estimated interaction $\widehat{\delta_{1;1}}$.

It is interesting that the spread of the estimated regression coefficients is smaller for the bootstrap than for the macroreplicates using LHS; nevertheless, the bootstrap gives more spread in the Pareto curves! The explanation is that the estimated regression coefficients in the metamodel for the standard deviation are negatively correlated (so they compensate variations in each other's values) in the macroreplicates, whereas they are independent in

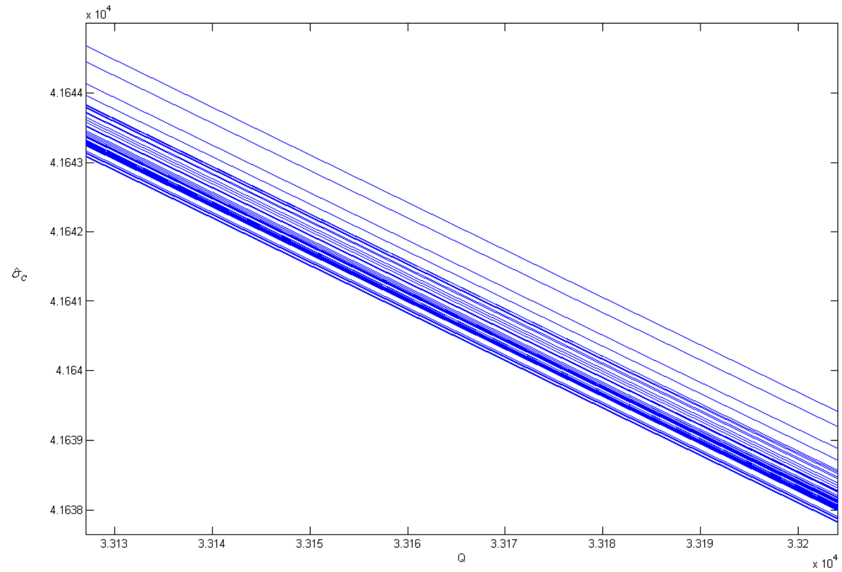

Fig. 12. Zoom: Standard deviation of Cost, estimated through 50 macroreplicates

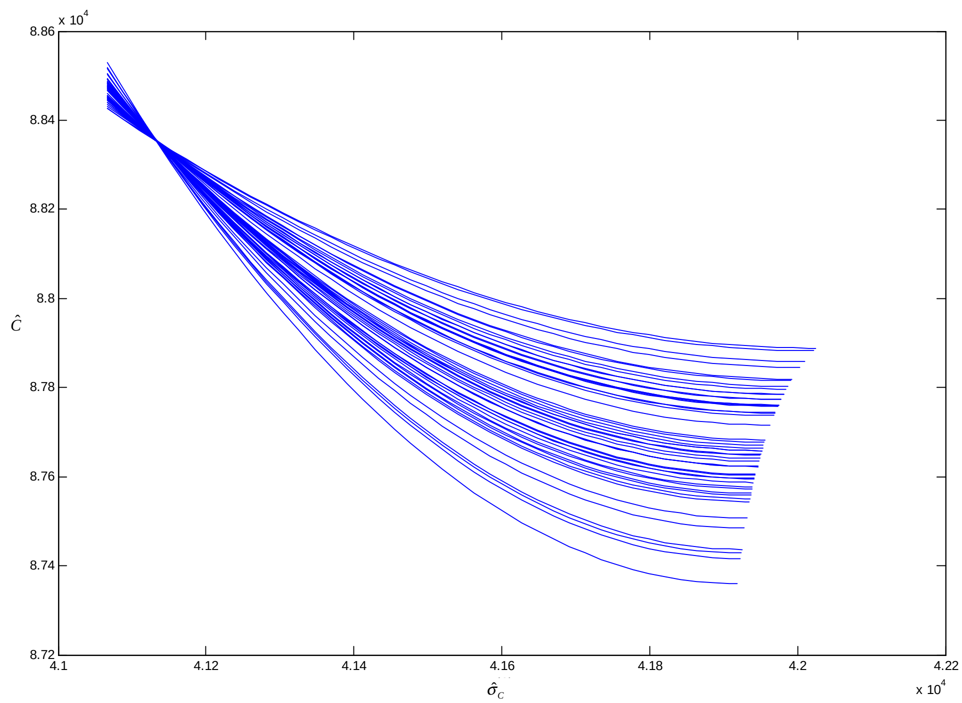

Fig. 13. Crude sampling: replicated estimated Pareto frontiers

the bootstrap. More precisely, this bootstrap uses a covariance matrix that implies $cov(\widehat{\gamma^*_{1(-i)}}, \widehat{\delta^*_{1;1(-i)}}) = 0$; for the macroreplicates we use Matlab's 'Symbolic Math Toolbox' to derive that the correlation coefficient $cor(\widehat{\gamma_{1(-i)}}, \widehat{\delta_{1;1(-i)}})$ is

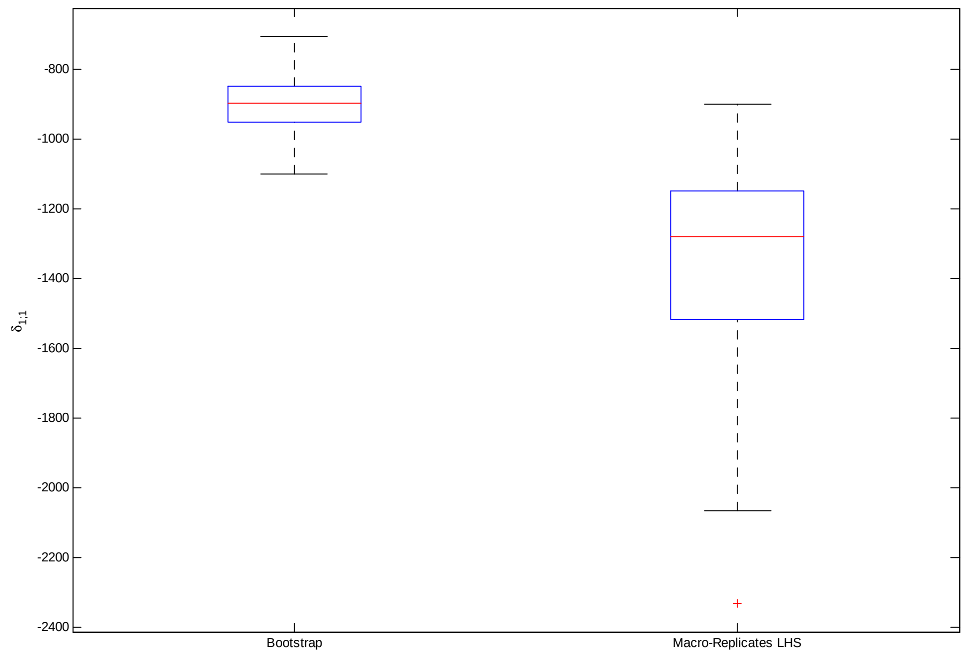

Fig. 14. Box plot for the estimated interaction $\delta_{1;1}$

-1; see Figure 15.

Appendix 4: CCD experiment

Table 7 and Figure 16 give details on our CCD experiment.

References

Alexandrov, N.M. and M.Y. Hussaini (1997), Multidisciplinary design optimization—state of the art. *Proceedings of the ICASE/NASA Langley Workshop on Multidisciplinary Design Optimization*, SIAM Proceedings Series

Apley, D.W., J. Liu, and W. Chen (2006), Understanding the effects of model uncertainty in robust design with computer experiments. *Journal of Mechanical Design*, 128, pp. 945-958

Ben-Tal, A. and A. Nemirovski (2008), Selected topics in robust convex optimization, *Mathematical Programming*, 112, no. 1, pp. 125-158

Beyer, H. and B. Sendhoff (2007), Robust optimization—a comprehensive survey. *Computer Methods in Applied Mechanics and Engineering,* 196, nos. 33-34, pp. 3190-3218

Chen, V.C.P., K.-L. Tsui, R.R. Barton, and J.K. Allen (2003), A review of

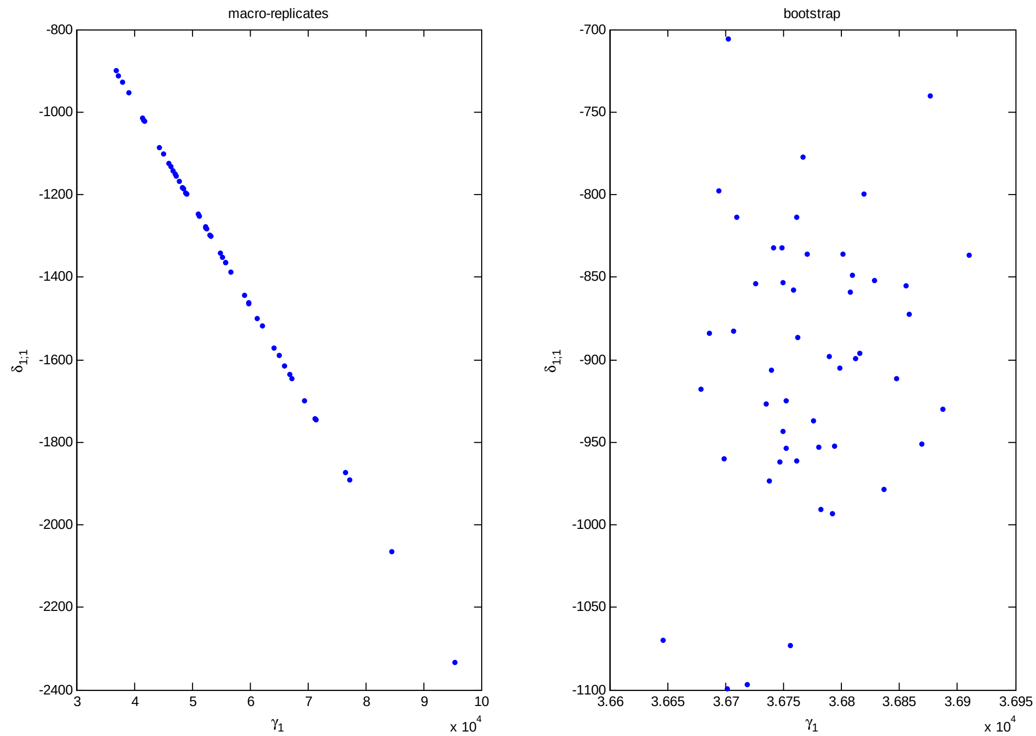

Fig. 15. Scatterplot of $\widehat{\gamma_1}$ and $\widehat{\delta_{1;1}}$ in macro-replicates (left-hand side) and bootstrap (right-hand side)

i	$\widehat{\beta_{0(-i)}}$	$\widehat{\beta_{1(-i)}}$	$\widehat{\beta_{1;1(-i)}}$	$\widehat{\gamma_{1(-i)}}$	$\widehat{\delta_{1;1(-i)}}$	$\widehat{y_{(-i)}}$	$\widehat{y_{(-i)}}/C_i$
0	88136.81	182.41	529.35	25915.14	-401.66		
1	88188.84	110.87	542.36	25843.59	-258.57	62518.17	1.0092
2	88155.43	156.81	534.00	25940.74	-452.86	114926.23	1.0018
3	88137.85	183.85	529.61	25913.70	-404.54	63342.16	1.0002
4	88104.44	137.91	521.26	25870.64	-490.66	114143.59	0.9969
5	88182.69	271.64	414.63	25915.14	-401.66	88627.80	0.9943
6	88110.59	233.40	594.89	25915.14	-401.66	89630.45	1.0032
7	88063.51	182.41	578.21	25962.65	-401.66	51346.78	0.9948
8	88178.26	182.41	501.71	25942.01	-401.66	124865.80	1.0012
9	88126.36	182.41	536.32	25915.14	-401.66	88126.36	0.9996

Table 7
Cross-validation of regression metamodel for RO of EOQ, based on CCD

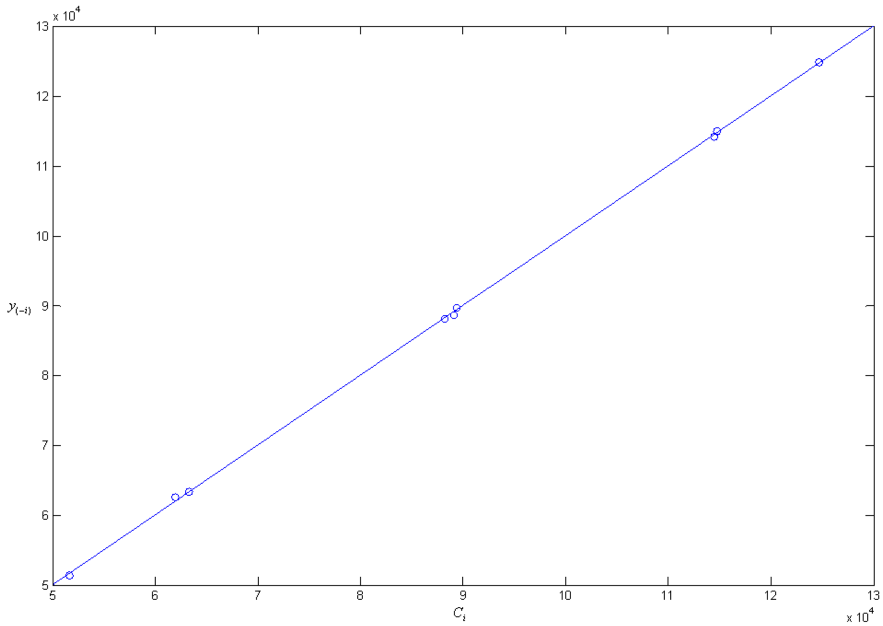

Fig. 16. Scatterplot of the EOQ regression metamodel for CCD

design and modeling in computer experiments. In: *Handbook of Statistics; Volume 22*, edited by R. Khattree and C.R. Rao, Elsevier, Amsterdam, pp. 231-261

Chen, V.C.P., K.-L. Tsui, R.R. Barton, and M. Meckesheimer (2006), A review on design, modeling, and applications of computer experiments. *IIE Transactions*, 38, pp. 273-291

Chen, W., R. Jin, and A. Sudjianto (2006), Analytical global sensitivity analysis and uncertainty propagation for robust design, *Journal of Quality Technology*, 38, no. 4, pp. 333-348

Chung, P.J., H.B. Goldfarb and D.C. Montgomery (2007), Optimal designs for mixture-process experiments with control and noise variables, *Journal of Quality Technology*, 39, no. 3, pp. 179-190

Darwish, M.A. (2008), EPQ models with varying setup cost. *International Journal of Production Economics*, 113, no. 1, pp. 297-306

Efron, B. and R.J. Tibshirani (1993), *An introduction to the bootstrap*. Chapman & Hall, New York

Greenberg, H.J. and T. Morisson (2008), Robust optimization. *Operations Research and Management Science Handbook*, edited by A.R. Ravindran, CRC Press, Boca Raton, Florida

Iooss, B., M. Ribatet, and A. Marrel (2007), Global sensitivity analysis of stochastic computer models with generalized additive models. Working Paper, CEA, Cadarache, Saint Paul lez Durance, France

Kelton, W.D., R.P. Sadowski, and D.T. Sturrock (2007), *Simulation with Arena; fourth edition*. McGraw-Hill, Boston

Kleijnen, J.P.C. (1980), *Computers and profits: quantifying financial benefits of information*. Addison-Wesley, Reading, Massachusetts

Kleijnen, J.P.C., (1987), *Statistical tools for simulation practitioners.* Marcel Dekker, New York

Kleijnen, J.P.C. (2008), *Design and analysis of simulation experiments*, Springer, New York

Kleijnen, J.P.C. and J.C. Helton (1999), Statistical analyses of scatter plots to identify important factors in large-scale simulations,1: review and comparison of techniques.*Reliability Engineering and Systems Safety*, 65, no. 2, pp. 147-185 (also published as Sandia Report SAND98-2202, April 1999)

Kleijnen, J.P.C. and R.G. Sargent (2000), A methodology for the fitting and validation of metamodels in simulation. *European Journal of Operational Research*, 120, no. 1, pp. 14-29

Marczyk, J. (2000), Stochastic multidisciplinary improvement: beyond optimization. 8th AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Long Beach, September 2000, AIAA-2000-4929

Meckesheimer, M., R.R. Barton, T.W. Simpson, and A.J. Booker (2001), Computationally inexpensive metamodel assessment strategies. *AIAA Journal*, 40. no. 10, pp. 2053-2060

Mula, J., R. Poler, J.P. García-Sabater, and F.C. Lario (2006), Models for production planning under uncertainty : A review. *International Journal of Production Economics*, 103, no. 1, pp. 271-285

Mulvey, J.M., R.J. Vanderbei, and S.A. Zenios (1995), Robust optimization of large-scale systems. *Operations Research*, 43, no. 2, pp. 264-281

Myers, R.H. and D.C. Montgomery (1995), *Response surface methodology: process and product optimization using designed experiments.* Wiley, New York

Oden, J.T., Chair (2006), *Revolutionizing engineering science through simulation.* National Science Foundation (NSF), Blue Ribbon Panel on Simulation-Based Engineering Science

Pentico, D.W., M.J. Drake, and C. Toews (2008), The deterministic EPQ with partial backordering: a new approach. *Omega*, in press

Sahinidis, N.V. (2004), Optimization under uncertainty: state-of-the-art and opportunities. *Computers and Chemical Engineering*, 28, pp. 971-983

Simpson, T.W., J. Peplinski, P.N. Koch, and J.K. Allen (2001), Metamodels for computer-based engineering design: survey and recommendation. *Engineering with Computers*, 17, no. 2, pp. 129-150

Stinstra, E. and D. den Hertog (2007), Robust optimization using computer experiments. *European Journal of Operational Research*, in press

Taguchi, G. (1987), *System of experimental designs, volumes 1 and 2.* UNIPUB/ Krauss International, White Plains, New York

Wu, C.F.J. and M. Hamada (2000), *Experiments; planning, analysis, and parameter design optimization.* Wiley, New York

Wu. J., J. Li, S. Wang and T.C.E. Cheng (2008), A note on mean-variance analysis of the newsvendor model with stockout cost *Omega*, in press

Yu, G. (1997), Robust economic order quantity models. *European Journal of Operational Research*, 100, no. 3, pp. 482-493